Partitioned Runge–Kutta–Chebyshev Methods for Diffusion-Advection-Reaction Problems

Zbinden, Christophe

How to cite

ZBINDEN, Christophe. Partitioned Runge–Kutta–Chebyshev Methods for Diffusion-Advection-Reaction Problems. In: SIAM journal on scientific computing, 2011, vol. 33, n° 4, p. 1707–1725. doi: 10.1137/100807892

This publication URL: https://archive-ouverte.unige.ch//unige:17070
Publication DOI: 10.1137/100807892

PARTITIONED RUNGE-KUTTA-CHEBYSHEV METHODS FOR DIFFUSION-ADVECTION-REACTION PROBLEMS

CHRISTOPHE J. ZBINDEN*

Abstract. An integration method based on RKC (Runge-Kutta-Chebyshev) methods is discussed which has been designed to treat moderately stiff and non-stiff terms separately. The method, called PRKC (Partitioned Runge-Kutta-Chebyshev), is a one-step, partitioned Runge-Kutta method of second-order. It belongs to the class of stabilized methods, *viz.* explicit Runge-Kutta methods possessing extended real stability intervals. The aim of the PRKC method is to reduce the number of function evaluations of the non-stiff terms and to get a non-zero imaginary stability boundary.

Key words. Numerical integration of differential equations, Runge-Kutta-Chebyshev methods, Stabilized second-order integration method, Partitioned Runge-Kutta methods

AMS subject classifications. 65L20, 65M12, 65M20

1. Introduction. Stabilized Runge-Kutta methods are explicit methods with extended stability domain along the negative real axis. Their real stability interval has a length proportional to the square of the number of stages. Therefore, these methods are especially suited for the time integration of moderately stiff systems, of large dimension and with eigenvalues known to lie in a long narrow strip along the negative axis. For instance, two- and three-dimensional parabolic partial differential equations (PDE) converted by the method of lines (MOL) give rise to this type of systems. Compared to implicit or IMEX (IMplicit EXplicit) methods, stabilized methods do not require the solution of large linear or nonlinear systems and have a low storage demand. Compared to general explicit methods, they avoid a too severe step size restriction.

There exist several stabilized methods. For example, the first- and second-order Runge-Kutta-Chebyshev (RKC) methods [19], the second- and fourth-order Orthogonal-Runge-Kutta-Chebyshev (ROCK) methods [4, 1], the third-order DUMKA method [10, 13, 11], and the second-order Stabilized Explicit Runge-Kutta (SERK2) methods [12]. To be able to treat very stiff reaction terms, [22] proposes to combine the RKC approach with the IMEX idea. A two-step method of this type is discussed in [17], which provides better stability on the imaginary axis. We also wish to mention the essentially optimal explicit Runge-Kutta methods [18] with stability domain containing a given set.

In this paper we discuss a one-step, stabilized method of second-order based on the RKC method. The method, called PRKC (Partitioned Runge-Kutta-Chebyshev), treats stiff and non-stiff terms separately. It is devoted to solve systems of ordinary differential equations (ODE) representing space discretization of PDEs as

$$\dot{Y}(t) = F\big(t, Y(t)\big) + G\big(t, Y(t)\big), \qquad t > 0, \qquad Y(t_0) = Y_0, \tag{1.1}$$

where $F : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ corresponds to diffusion terms and $G : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ to advection and/or non-stiff reaction terms. The F-function is discretized by a RKC method which typically has a large number of internal stages. In contrast, only 4 function calls of G are needed for each step. We have chosen RKC rather than

*Section de Mathématiques, Université de Genève, 2-4 rue du Lièvre, CH-1211 Genève 4, Switzerland. (Christophe.Zbinden@unige.ch).

another method because all its coefficients are available in analytical form for an arbitrary number of stages.

The paper is organized as follows. In Section 2 we will give a short review about the RKC methods. We will introduce the PRKC method and discuss its order of consistency in Section 3. Semi-discretizations of hyperbolic-parabolic PDEs are a relevant application. These discretizations exhibit eigenvalues in a narrow strip along the negative axis of the complex plane. As a model, we will investigate linear advection-diffusion problems with our method in Section 4. In Section 5, the actual implementation of the second-order method into a FORTRAN code `PRKC` will be discussed, in particular the mechanisms for the step length selection and the error estimate. The last section of this paper is devoted to numerical results obtained with `PRKC`.

2. The explicit RKC method. The RKC methods are stabilized methods dedicated to solve initial value problems for ODEs of the form

$$\dot{Y}(t) = F\big(t, Y(t)\big), \qquad t > 0, \qquad Y(t_0) = Y_0, \tag{2.1}$$

where the Jacobian matrix $\frac{\partial F}{\partial Y}(t, Y)$ has all its eigenvalues near the negative real axis.

Let Y_n denote the approximation to $Y(t)$ at $t = t_n$ and let $\tau = t_{n+1} - t_n$ be the step size in the current step from t_n to t_{n+1}. The explicit m-stage RKC formula of second-order is of the form [8, Section V.1]

$$\begin{aligned}
K_0 &= Y_n, \\
K_1 &= K_0 + \kappa_1 \tau F_0, \\
K_j &= (1 - \mu_j - \nu_j) K_0 + \mu_j K_{j-1} + \nu_j K_{j-2} + \kappa_j \tau F_{j-1} - a_{j-1} \kappa_j \tau F_0, \\
&\qquad j = 2, \ldots, m, \\
Y_{n+1} &= K_m.
\end{aligned} \tag{2.2}$$

The coefficients are defined as follows

$$\omega_0 = 1 + \frac{\eta}{m^2}, \qquad \omega_1 = \frac{T'_m(\omega_0)}{T''_m(\omega_0)}, \tag{2.3}$$

where $\eta \geq 0$ is a damping parameter (usually set to 2/13), and T_m are the Chebyshev polynomials of the first kind defined in (2.6),

$$\begin{aligned}
b_j &= \frac{T''_j(\omega_0)}{\big(T'_j(\omega_0)\big)^2}, \ j = 2, \ldots, m, \quad b_0 = b_2, \quad b_1 = b_2, \\
a_j &= 1 - b_j T_j(\omega_0), \ j = 0, \ldots, m, \\
\mu_j &= \frac{2 b_j \omega_0}{b_{j-1}}, \quad \nu_j = \frac{-b_j}{b_{j-2}}, \quad \kappa_1 = b_1 \omega_1, \quad \kappa_j = \frac{2 b_j \omega_1}{b_{j-1}}, \ j = 2, \ldots, m.
\end{aligned} \tag{2.4}$$

In (2.2), F_j denotes $F\left(t_n + c_j \tau,\ K_j\right)$ where c_j are

$$c_j = \omega_1 \frac{T''_j(\omega_0)}{T'_j(\omega_0)}, \ j = 2, \ldots, m, \quad c_1 = \frac{c_2}{T'_2(\omega_0)}, \quad c_0 = 0. \tag{2.5}$$

Note that $0 = c_0 < c_1 < \ldots < c_{m-1} < c_m = 1$, and thus all values $t_n + c_j \tau$ lie within the current integration step.

Due to the specific recursive nature of the method coming from the Chebyshev polynomials of the first kind $T_m(z)$, $z \in \mathbb{C}$

$$T_j(z) = 2zT_{j-1}(z) - T_{j-2}(z),\ j = 2, \ldots, m, \quad T_1(z) = z, \quad T_0(z) = 1, \tag{2.6}$$

formula (2.2) is more convenient to work with than the common RK formula.

Applied to the scalar stability test equation

$$\dot{y} = \lambda y, \quad y_0 = 1, \quad z = \tau\lambda \in \mathbb{C}, \tag{2.7}$$

the RKC method gives for the internal stages

$$R_j(z) = a_j + b_j T_j(\omega_0 + \omega_1 z), \qquad j = 0, \ldots, m. \tag{2.8}$$

Using (2.5), we see that $R_j(z)$ approximates $e^{c_j z}$ for $z \to 0$ with second order accuracy, except for the first-stage formula which is necessarily of first order[1].

The polynomial $R_m(z)$ of degree m in z in (2.8) is the stability function. Its real stability boundary $\beta_{\mathbb{R}}(m)$, *i.e.* the maximum number such that all the points $x \in [-\beta_{\mathbb{R}}(m),\ 0]$ lie in the stability region $\mathcal{S} = \{z \in \mathbb{C};\ |R_m(z)| \leq 1\}$, can be seen to satisfy [21]

$$\beta_{\mathbb{R}}(m) \approx \frac{\omega_0 + 1}{\omega_1} \approx \frac{2}{3}(m^2 - 1)\Big(1 - \frac{2}{15}\eta\Big). \tag{2.9}$$

If the damping is strictly positive, $\eta > 0$, then the stability function satisfies $0 < R_m(x) < 1$ in the interior of the real stability interval $[-\beta_{\mathbb{R}}(m),\ 0]$, but $\beta_{\mathbb{R}}(m)$ becomes slightly smaller with increasing η.

3. The explicit partitioned RKC method. In this section, we introduce a new class of stabilized methods for the solution of moderately stiff ODEs written in the form (1.1). The aim of these methods is to reduce the number of evaluations of the non-stiff terms while keeping a large stability domain.

Inspired by the S-ROCK (stochastic orthogonal Runge-Kutta-Chebyshev) [2, 3] concept, we define for arbitrary $m \geq 2$ the m-stage[2] explicit partitioned RKC method by

$$\begin{aligned}
K_{-1} &= Y_n,\\
K_0 &= K_{-1} + \alpha_0\tau G_{-1},\\
K_1 &= K_0 + \kappa_1\tau F_0,\\
K_j &= (1-\mu_j-\nu_j)K_0 + \mu_j K_{j-1} + \nu_j K_{j-2} + \kappa_j\tau F_{j-1} - a_{j-1}\kappa_j\tau F_0,\\
&\quad\ j = 2, \ldots, m-1,\\
K_m &= (1-\mu_m-\nu_m)K_0 + \mu_m K_{m-1} + \nu_m K_{m-2} + \kappa_m\tau F_{m-1} - a_{m-1}\kappa_m\tau F_0\\
&\quad +\alpha_1\tau G_{-1} + \alpha_2\tau G_0 + \alpha_3\tau G_{m-1},\\
K_{m+1} &= (1-\mu_m-\nu_m)K_0 + \mu_m K_{m-1} + \nu_m K_{m-2} + \kappa_m\tau F_{m-1} - a_{m-1}\kappa_m\tau F_0\\
&\quad +\alpha_4\tau G_{-1} + \alpha_5\tau G_0 + \alpha_6\tau G_{m-1} + \alpha_7\tau G_m,\\
Y_{n+1} &= K_{m+1},
\end{aligned} \tag{3.1}$$

[1] In the original method [19] all intermediate approximations are of order one. Here we adopt the choice of parameters made in [16].

[2] m indicates the number of stages of the RKC method.

where $\{\alpha_i\} \subset \mathbb{R}$ will be defined later, $G_{-1} = G(t_n, K_{-1})$, $G_0 = G(t_n + \alpha_0\tau, K_0)$, $G_{m-1} = G(t_n + \alpha_0\tau, K_{m-1})$, $G_m = G(t_n + \tau, K_m)$ and all other coefficients are identical to the RKC method (2.3), (2.4) and (2.5). We observe that for $G \equiv 0$ this method is identical to the RKC method (2.2).

Note that the first stages look like a splitting using the first-order forward Euler and the second-order RKC methods. However the whole method (3.1) does not belong to the class of splitting methods, because the final stages use information of several internal stages. These last two stages provide greater freedom.

The method is a partitioned RK method of the form (for notational convenience we suppress the dependence on t for F and G)

$$\widetilde{K}_i = Y_n + \tau\Big(\sum_{j=1}^{i-1} \mathsf{a}_{ij} F(\widetilde{K}_{j-2}) + \sum_{j=1}^{i-1} \widehat{\mathsf{a}}_{ij} G(\widetilde{K}_{j-2})\Big), \qquad i = 1, \dots, m+3. \tag{3.2}$$

The corresponding Butcher tableau[3] for the F-method is

$$\begin{array}{c|cccccccc}
\mathsf{c}_1 & \mathsf{a}_{11}=0 & & & & & & & \\
\mathsf{c}_2 & \mathsf{a}_{21}=0 & \mathsf{a}_{22}=0 & & & & & & \\
\mathsf{c}_3 & 0 & * & 0 & & & & & \\
\mathsf{c}_4 & 0 & * & * & 0 & & & & \\
\vdots & \vdots & \vdots & \vdots & \ddots & 0 & & & \\
\mathsf{c}_{m+2} & 0 & \mathsf{a}_{m+2,2}=* & \dots & & * & 0 & & \\
\mathsf{c}_{m+3} & 0 & \mathsf{a}_{m+2,2} & \dots & & \mathsf{a}_{m+2,m+1} & 0 & 0 & \\
\hline
 & \mathsf{b}_1=0 & \mathsf{a}_{m+2,2} & \dots & & \mathsf{a}_{m+2,m+1} & \mathsf{b}_{m+2}=0 & \mathsf{b}_{m+3}=0
\end{array} \tag{3.3}$$

where "$*$" represents a non zero entry, and for the G-method it is

$$\begin{array}{c|cccccccc}
\widehat{\mathsf{c}}_1 & \widehat{\mathsf{a}}_{11}=0 & & & & & & & \\
\widehat{\mathsf{c}}_2 & \widehat{\mathsf{a}}_{21}=\alpha_0 & \widehat{\mathsf{a}}_{22}=0 & & & & & & \\
\widehat{\mathsf{c}}_3 & \alpha_0 & 0 & 0 & & & & & \\
\widehat{\mathsf{c}}_4 & \alpha_0 & 0 & \dots & 0 & & & & \\
\vdots & \vdots & \vdots & & & \ddots & & & \\
\widehat{\mathsf{c}}_{m+1} & \alpha_0 & 0 & & \dots & & 0 & & \\
\widehat{\mathsf{c}}_{m+2} & \alpha_0+\alpha_1 & \alpha_2 & 0 & \dots & 0 & \alpha_3 & 0 & \\
\widehat{\mathsf{c}}_{m+3} & \alpha_0+\alpha_4 & \alpha_5 & 0 & \dots & 0 & \alpha_6 & \alpha_7 & 0 \\
\hline
 & \widehat{\mathsf{b}}_1=\alpha_1+\alpha_4 & \alpha_4 & 0 & \dots & 0 & \alpha_6 & \widehat{\mathsf{b}}_{m+2}=\alpha_7 & \widehat{\mathsf{b}}_{m+3}=0
\end{array} \tag{3.4}$$

with $\widehat{\mathsf{c}}_i = \sum \widehat{\mathsf{a}}_{ij}$. In the array (3.3)

- the last line $\{\mathsf{a}_{m+3,k}\}$ and weights $\{\mathsf{b}_k\}$ are a copy of the second last line of the $\{\mathsf{a}_{ij}\}$ matrix,
- the sub matrix $\{\mathsf{a}_{ij}\}_{2\le i,j\le m+2}$ is the Butcher tableau of the RKC method (2.2), we have $\mathsf{c}_1 = 0$, $\mathsf{c}_{j+2} = c_j$ for $j = 0, \dots, m$ and $\mathsf{c}_{m+3} = 1$.

We observe that the number of F-evaluations (stiff terms) per step depends on the stage number. The array (3.4) shows that the method only needs four G-evaluations (non-stiff terms).

[3] Caution, sans serif font (a, b and c) indicates Butcher's array coefficients, not RKC coefficients.

3.1. Discussion of order. We shall show here how to construct a family of second order methods of the type (3.1). Configurations with a G-method of order 3 are designed to achieve better stability on the imaginary axes (see Sect. 4.1).

THEOREM 3.1. *The partitioned Runge-Kutta-Chebyshev method* (3.1) *is of order* 2*, if*

$$\begin{gathered}\alpha_0+\alpha_4+\alpha_5+\alpha_6+\alpha_7=1,\quad \alpha_0(\alpha_5+\alpha_6)+\alpha_7(\alpha_0+\alpha_1+\alpha_2+\alpha_3)=1/2,\\ \alpha_0=1/2,\qquad\qquad c_{m-1}\alpha_6+\alpha_7=1/2.\end{gathered}\tag{3.5}$$

In addition, the G-method is of order 3 *if*

$$\begin{gathered}\alpha_0=1/2,\qquad \alpha_1=-1/2+v(3-4v),\qquad \alpha_2=2v(2v-1)-\alpha_3,\\ \alpha_4=\frac{1-3v}{6v},\ \alpha_5=\frac{1+3v(1-2v)+4c_{m-1}v(3v-2)}{6c_{m-1}v(2v-1)},\ \alpha_6=\frac{3v(2v-1)-1}{6c_{m-1}v(2v-1)},\\ \alpha_7=\frac{1}{6v(2v-1)}.\end{gathered}\tag{3.6}$$

where $v\notin\{0,\,1/2\}$ and $\alpha_3\in\mathbb{R}$ are free parameters.

Proof. We know from [6, p. 308] that a partitioned method like (3.2) is of order 2, if and only if

1. each of the two Runge-Kutta schemes has order 2
2. and the coupling conditions

$$\sum_i \mathsf{b}_i\widehat{\mathsf{c}}_i=\frac{1}{2},\qquad\qquad \sum_i \widehat{\mathsf{b}}_i\mathsf{c}_i=\frac{1}{2},\tag{3.7}$$

are satisfied.

We already know that the F-method (3.3) is of order 2, because 2 is the order of the RKC method. For the G-method we consider the case $F\equiv 0$. Only three evaluations of the G-function are distinct in this case, the array (3.4) becomes the 3-stage RK method

$$\begin{array}{c|ccc} 0 & & & \\ u & \alpha_0 & & \\ v & \alpha_0+\alpha_1 & \alpha_2+\alpha_3 & \\ \hline & \alpha_0+\alpha_4 & \alpha_5+\alpha_6 & \alpha_7 \end{array}\tag{3.8}$$

where $u=\alpha_0$ and $v=\sum_{i=0}^{3}\alpha_i$. The conditions for order 2 of RK methods are $\sum \mathsf{b}_i=1$ and $\sum \mathsf{b}_i\mathsf{c}_i=1/2$. Applied to the array (3.8) they give rise to the first line of (3.5). Using (3.3) and (3.4) the coupling conditions (3.7) give the second line of (3.5), which completes the proof of the first part of the theorem.

The second part is obtained by comparing the array (3.8) with the general 3-stage RK method of order 3 (see [6, p. 142])

$$\begin{array}{c|ccc} 0 & & & \\ u & u & & \\ v & v-\dfrac{v(v-u)}{u(2-3u)} & \dfrac{v(v-u)}{u(2-3u)} & \\ \hline & 1-\mathsf{b}_2-\mathsf{b}_3 & \dfrac{2-3v}{6u(u-v)} & \dfrac{2-3u}{6v(v-u)} \end{array}$$

□

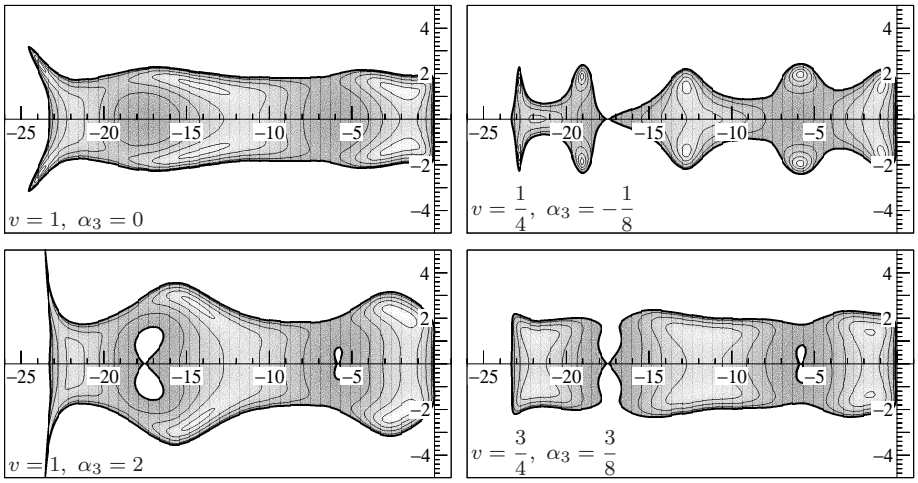

FIG. 3.1. *Stability regions in the* $(\tau\lambda, \tau\mu)$*-space for the 6-stage undamped PRKC method,* $\eta = 0$.

3.2. Stability region. Following [5, 9], we adapt (2.7) to our method by defining the modified test equation

$$\dot{y} = \lambda y + \mathrm{i}\mu y, \quad y_0 = 1, \quad \lambda, \mu \in \mathbb{R}. \tag{3.9}$$

It gives information for systems $\dot{y} = Ay + By$ where A and B can be diagonalized with the same transformation T and the eigenvalues of A are real, and those of B on the imaginary axis. It is especially suited for advection-diffusion equations with space discretized by second-order central differences.

Applied to the test equation (3.9), method (3.1) gives for the internal stages $\widetilde{R}_j$

$$\begin{aligned}
\widetilde{R}_{-1}(\tau\lambda, \tau\mu) &= 1,\\
\widetilde{R}_0(\tau\lambda, \tau\mu) &= 1 + \alpha_0\tau\mathrm{i}\mu,\\
\widetilde{R}_j(\tau\lambda, \tau\mu) &= R_j(\tau\lambda)(1 + \alpha_0\tau\mathrm{i}\mu), \quad j = 1, \ldots, m-1,\\
\widetilde{R}_m(\tau\lambda, \tau\mu) &= R_m(\tau\lambda)(1 + \alpha_0\tau\mathrm{i}\mu) + (\alpha_1 + \alpha_2)\tau\mathrm{i}\mu + \alpha_0\alpha_2(\tau\mathrm{i}\mu)^2\\
&\quad + R_{m-1}(\tau\lambda)\big(\alpha_3\tau\mathrm{i}\mu + \alpha_0\alpha_3(\tau\mathrm{i}\mu)^2\big),\\
\widetilde{R}_{m+1}(\tau\lambda, \tau\mu) &= R_m(\tau\lambda)\big(1 + (\alpha_0 + \alpha_7)\tau\mathrm{i}\mu + \alpha_0\alpha_7(\tau\mathrm{i}\mu)^2\big)\\
&\quad + R_{m-1}(\tau\lambda)\big(\alpha_6\tau\mathrm{i}\mu + (\alpha_0\alpha_6 + \alpha_3\alpha_7)(\tau\mathrm{i}\mu)^2 + \alpha_0\alpha_3\alpha_7(\tau\mathrm{i}\mu)^3\big)\\
&\quad + (\alpha_4 + \alpha_5)\tau\mathrm{i}\mu + \big(\alpha_0\alpha_5 + (\alpha_1 + \alpha_2)\alpha_7\big)(\tau\mathrm{i}\mu)^2 + \alpha_0\alpha_2\alpha_7(\tau\mathrm{i}\mu)^3,
\end{aligned} \tag{3.10}$$

where $R_j(\tau\lambda)$ denotes $a_j + b_j T_j(\omega_0 + \omega_1\tau\lambda)$ as in (2.8).

The stability function is given by the last internal stage $\widetilde{R}_{m+1}(\tau\lambda, \tau\mu)$ in (3.10) and the stability region $\mathcal{S}$ is defined by

$$\mathcal{S} := \big\{(x, y) \in \mathbb{R}^2;\ \big|\widetilde{R}_{m+1}(x, y)\big| \le 1\big\}. \tag{3.11}$$

Notice that in contrast to the stability function (2.8) of RKC methods, the function $\widetilde{R}_{m+1}(x, y)$ in (3.10) is not an analytic function of $z = x + \mathrm{i}y$ which is a consequence of the use of the modified test equation (3.9).

FIG 3.1 and FIG 3.2 show the stability regions of the PRKC method for various values of v and α_3 in (3.6). The imaginary stability boundary $\beta_{\mathbb{I}}$, *i.e.* the maximum

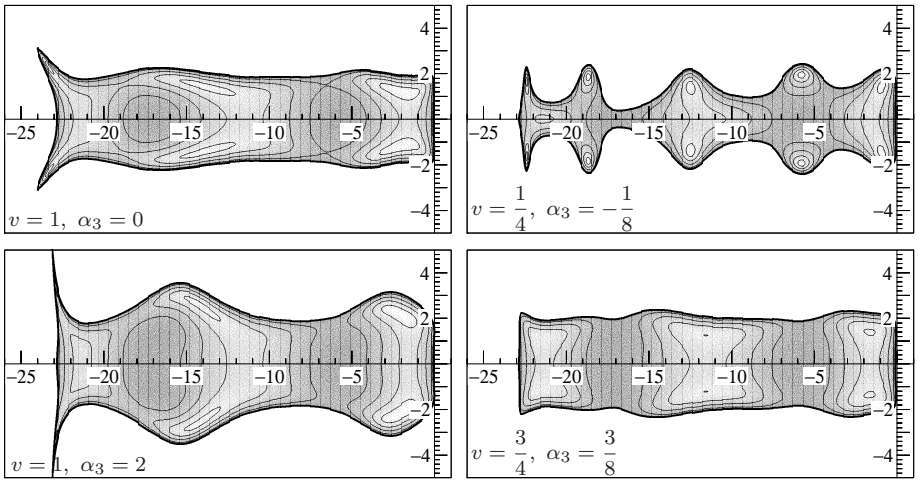

FIG. 3.2. *Stability regions in the* $(\tau\lambda,\, \tau\mu)$*-space for the 6-stage damped PRKC method,* $\eta = 2/13$.

number such that all the points of the imaginary axis $y \in [-\mathrm{i}\beta_{\mathbb{I}},\, \mathrm{i}\beta_{\mathbb{I}}]$ lie in the stability region, is equal to $\sqrt{3}$ for all values of v and α_3. Note also that even without damping the stability region possesses a strip around the negative real axis for certain values of v and α_3.

Consider first the undamped case $\eta = 0$. Using $T_j(1) = 1$, $T_j'(1) = j^2$ and $T_j''(1) = \frac{1}{3}j^2(j^2-1)$, the stability function for $m \geq 3$ is given by substituting

$$R_{m-1}(\tau\lambda) = \frac{2}{3} + \frac{1}{3(m-1)^2} + \Big(\frac{1}{3} - \frac{1}{3(m-1)^2}\Big)T_{m-1}\Big(1 + \frac{3\tau\lambda}{m^2-1}\Big),$$
$$R_m(\tau\lambda) = \frac{2}{3} + \frac{1}{3m^2} + \Big(\frac{1}{3} - \frac{1}{3m^2}\Big)T_m\Big(1 + \frac{3\tau\lambda}{m^2-1}\Big), \tag{3.12}$$

into $\widetilde{R}_{m+1}(\tau\lambda,\, \tau\mu)$ in (3.10).

THEOREM 3.2. *If* $v = 1$ *and* $\alpha_3 \leq 0$ *in* (3.6)*, then for* $m \geq 2$ *the stability region of the method* (3.1) *without damping possesses a strip around the negative real axis, i.e. there exist positive constants* $k_1 > 0,\ k_2 > 0$ *such that*

$$\big\{(x,\, y) \in \mathbb{R}^2;\ -k_1 m^2 \leq x \leq 0,\ |y| \leq k_2\big\} \subset \mathcal{S}.$$

Proof. We shall use a geometric argument to show that it suffices to treat only four cases of pairs $\big(R_m(\tau\lambda),\, R_{m-1}(\tau\lambda)\big)$. To simplify the notation, we omit the arguments in the stability function (3.10) and write it as

$$\widetilde{R}_{m+1} = R_m A + R_{m-1} B + C$$

where $R_m,\ R_{m-1}$ are the real polynomials (2.8) with real argument $\tau\lambda$ and

$$A = 1 + \frac{2}{3}\tau\mathrm{i}\mu + \frac{1}{12}(\tau\mathrm{i}\mu)^2,$$
$$B = \frac{1}{3c_{m-1}}\tau\mathrm{i}\mu + \Big(\frac{\alpha_3}{6} + \frac{1}{6c_{m-1}}\Big)(\tau\mathrm{i}\mu)^2 + \frac{\alpha_3}{12}(\tau\mathrm{i}\mu)^3,$$
$$C = \Big(\frac{1}{3} - \frac{1}{3c_{m-1}}\Big)\tau\mathrm{i}\mu + \Big(\frac{5-2\alpha_3}{12} - \frac{1}{6c_{m-1}}\Big)(\tau\mathrm{i}\mu)^2 + \Big(\frac{1}{6} - \frac{\alpha_3}{12}\Big)(\tau\mathrm{i}\mu)^3,$$

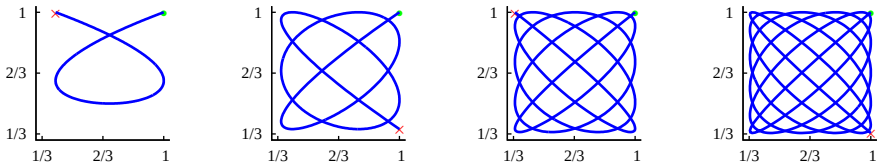

FIG. 3.3. $R_{m-1}(\tau\lambda)$ *vs.* $R_m(\tau\lambda)$ *for* $-\beta_{\mathbb{R}}(m)$ (×) $\leq \tau\lambda \leq 0$ (●) *and* $m = 3,\ 6,\ 9,\ 12$.

are real polynomials with complex argument $\tau i\mu$. We have

$$\left|\widetilde{R}_{m+1}\right|^2 = (R_m,\ R_{m-1}) \begin{pmatrix} |A|^2 & \Re\left(A\bar{B}\right) \\ \Re\left(A\bar{B}\right) & |B|^2 \end{pmatrix} \begin{pmatrix} R_m \\ R_{m-1} \end{pmatrix} + \left(2\Re\left(A\bar{C}\right),\ 2\Re\left(B\bar{C}\right)\right) \begin{pmatrix} R_m \\ R_{m-1} \end{pmatrix} + |C|^2 . \tag{3.13}$$

Recall that R_{m-1} and R_m map points within $[-\beta_{\mathbb{R}}(m),\ 0]$ into the interval $[1/3,\ 1]$ (see (3.12)), so that the curve of R_{m-1} *versus* R_m stays in the square $[1/3,\ 1]^2$ (illustrated in FIG 3.3)[4]. By freezing $\tau\mu$ and m, the polynomials A, B and C become constant and (3.13), considered as a function of $(R_m,\ R_{m-1})$, is a convex quadratic function. Therefore the maximum over the square $[1/3,\ 1]^2$ is necessarily taken in one of its corners. Note that cases representing the corners of the square can occur when T_{m-1} and T_m in (3.12) are in phase or in opposite phase.

For each case, either we calculate the values of $\tau\mu$ for which $\left|\widetilde{R}_{m+1}\right|^2 \leq 1$, or we give a condition on α_3 to get a strip around the negative real axis.

Case $R_m = R_{m-1} = 1$: We have

$$\left|\widetilde{R}_{m+1}\right|^2 = \frac{1}{36}(\tau\mu)^6 - \frac{1}{12}(\tau\mu)^4 + 1 \leq 1 \quad \text{for} \quad \tau\mu \in \left[-\sqrt{3},\ \sqrt{3}\right]. \tag{3.14}$$

Case $R_m = 1/3,\ R_{m-1} = 1$: We have

$$\left|\widetilde{R}_{m+1}\right|^2 = \frac{1}{36}(\tau\mu)^6 + \frac{1}{81}(\tau\mu)^4 + \frac{1}{81}(\tau\mu)^2 + \frac{1}{9} \leq 1 \quad \text{for} \quad \tau\mu \in [-1.7288,\ 1.7288]\,.$$

Case $R_m = R_{m-1} = 1/3$: We have

$$\left|\widetilde{R}_{m+1}\right|^2 = \left(\frac{1}{324}\,{\alpha_3}^2 - \frac{1}{54}\,\alpha_3 + \frac{1}{36}\right)(\tau\mu)^6 + \left(\frac{1}{81}\,{\alpha_3}^2 - \frac{1}{27}\,\alpha_3 + \frac{1}{81} - \frac{2}{81\,c_{m-1}} + \frac{1}{81\,{c_{m-1}}^2}\right)(\tau\mu)^4 + \left(\frac{2}{27}\,\alpha_3 + \frac{1}{81} - \frac{14}{81\,c_{m-1}} + \frac{4}{81\,{c_{m-1}}^2}\right)(\tau\mu)^2 + \frac{1}{9}.$$

For small values of $\tau\mu$, neglecting the terms with $(\tau\mu)^6$ and $(\tau\mu)^4$, the expression can be bounded for $c_{m-1} \geq 2/7$ (which is true for $m \geq 3$, see (2.5); the case $m = 2$ follows from the remark below) by

$$\left(\frac{2}{27}\,\alpha_3 + \frac{1}{81}\right)(\tau\mu)^2 + \frac{1}{9}.$$

[4]For large m, the curve of R_{m-1} *versus* R_m approaches the corners of the square $[1/3,\ 1]^2$

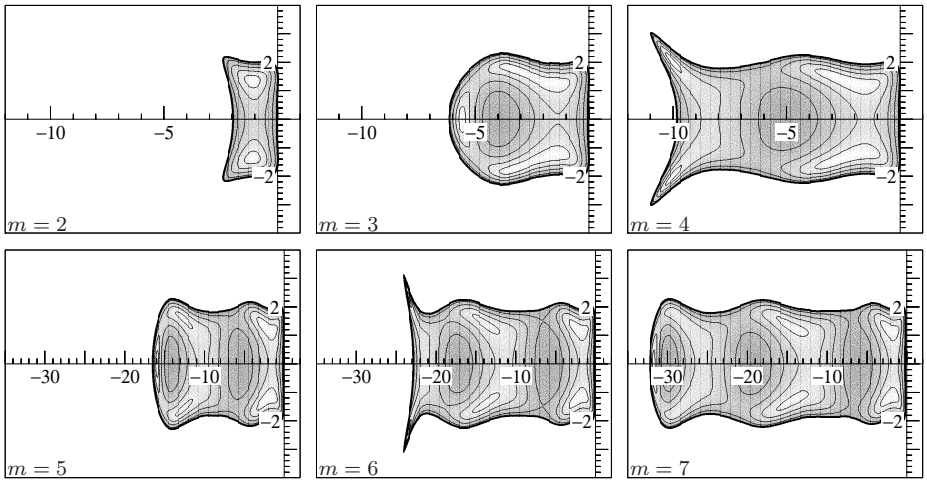

FIG. 3.4. *Stability regions in the* $(\tau\lambda, \tau\mu)$*-space of PRKC method with* $v = 1$, $\alpha_3 = 0$ *and* $\eta = 2/13$ *for* $m = 2, 3, \ldots, 7$.

Because of the constant term, this case does not impose severe restrictions on α_3 to get a thin strip around the negative real axis.
Case $R_m = 1,\ R_{m-1} = 1/3$: We have

$$\begin{aligned}\left|\widetilde{R}_{m+1}\right|^2 = &\left(\frac{1}{324}\,{\alpha_3}^2 - \frac{1}{54}\,\alpha_3 + \frac{1}{36}\right)(\tau\mu)^6 + \\ &\left(\frac{1}{81}\,{\alpha_3}^2 - \frac{1}{12} - \frac{1}{27\,c_{m-1}} + \frac{1}{81\,{c_{m-1}}^2}\right)(\tau\mu)^4 + \\ &\left(\frac{2}{9}\,\alpha_3 - \frac{2}{9\,c_{m-1}} + \frac{4}{81\,{c_{m-1}}^2}\right)(\tau\mu)^2 + 1.\end{aligned}$$

As in the previous case, by using $c_{m-1} \geq 0.25$ for $m \geq 2$, we get

$$\frac{2}{9}\alpha_3(\tau\mu)^2 + 1,$$

imposing $\alpha_3 \leq 0$. □

Remark: For the case $m = 2$ and arbitrary damping η, ω_1 is equal to ω_0. Moreover, ω_0 does not appear in $R_m(\tau\lambda) = 1 + \tau\lambda + \frac{1}{2}(\tau\lambda)^2$, but only in $R_{m-1}(\tau\lambda) = 1 + \frac{1}{4\omega_0}\tau\lambda$. Hence, the pair $(R_m,\, R_{m-1})$ stays in the square $[1/2,\, 1]^2$, whenever the argument is in $[-\beta_{\mathbb{R}}(2),\, 0] = [-2,\, 0]$. This implies that we have to consider the case $R_m = R_{m-1} = 1/2$ instead of $R_m = R_{m-1} = 1/3$ in the previous proof.

By continuity, the existence of such a strip can be extended to small values of damping $1 \gg \eta > 0$. Notice that a similar study can be done for other values of v.

3.3. Choice of coefficients. In the rest of this article, PRKC denotes the method (3.1) with parameters $v = 1$ and $\alpha_3 = 0$ in (3.6) and damping $\eta = 2/13$ (as in [21]). This choice of damping reduces the real stability boundary $\beta_{\mathbb{R}}(m)$ of about 2% compared to the undamped case ($\eta = 0$). Note that for $v = 1$, the under-

lying quadrature rule corresponding to the Butcher tableau (3.8)

$$\begin{array}{c|ccc} 0 & & & \\ 1/2 & 1/2 & & \\ 1 & -1 & 2 & \\ \hline & 1/6 & 4/6 & 1/6 \end{array} \tag{3.15}$$

is Simpson's rule. Moreover, since $\alpha_2 + \alpha_3 = 2$ (third line of (3.15)) and $\alpha_3 \le 0$ (Theorem 3.2), we chose $\alpha_3 = 0$ to avoid too large coefficients in the second last stage of the PRKC method (3.1).

Applied to the test equation (3.9), PRKC gives the stability function

$$\begin{aligned} \widetilde{R}_{m+1}(\tau\lambda,\, \tau\mu) = R_m(\tau\lambda)\Big(1 + \frac{2}{3}\tau\mathrm{i}\mu + \frac{1}{12}(\tau\mathrm{i}\mu)^2\Big) + \\ R_{m-1}(\tau\lambda)\Big(\frac{1}{3c_{m-1}}\tau\mathrm{i}\mu + \frac{1}{6c_{m-1}}(\tau\mathrm{i}\mu)^2\Big) + \\ \Big(\frac{1}{3} - \frac{1}{3c_{m-1}}\Big)\tau\mathrm{i}\mu + \Big(\frac{5}{12} - \frac{1}{6c_{m-1}}\Big)(\tau\mathrm{i}\mu)^2 + \frac{1}{6}(\tau\mathrm{i}\mu)^3. \end{aligned} \tag{3.16}$$

FIG 3.4 shows stability regions of PRKC method for several number of stages m.

THEOREM 3.3. *Consider the method* (3.1) *with* $v = 1$, $\alpha_3 = 0$ *and damping parameter* $\eta = 2/13$. *For* $m \ge 2$, *its stability region contains the rectangle*

$$\mathcal{R}(m) := \big\{(x,\, y) \in \mathbb{R}^2;\ -\beta_{\mathbb{R}}(m) \le x \le 0,\ |y| \le 1.7273\big\} \tag{3.17}$$

where $\beta_{\mathbb{R}}(m) = 0.65(m^2 - 1)$ *as in* (2.9), *see also Table 3.1.*

Proof. We shall follow the outline of the proof of the Theorem 3.2. Since R_{m-1}, R_m and c_{m-1} in the stability function (3.16) depend on η, which is now fixed to $2/13$, we first need to investigate the range of R_{m-1} and R_m for the domain $[-\beta_{\mathbb{R}}(m),\ 0]$ to determine the four cases to study.

Recall that for $x \in \mathbb{R}$ and $2 \le j \le m$ we have $R_j(x) = a_j + b_j T_j(\omega_0 + \omega_1 x)$, where ω_0, ω_1 are defined in (2.3) and a_j, b_j in (2.4). For $-\beta_{\mathbb{R}}(m) \le x \le 0$, by using $-1 \le T_j(\omega_0 + \omega_1 x) \le T_j(\omega_0)$, we can bound $R_j(x)$ by

$$a_j - b_j \le R_j(x) \le a_j + b_j T_j(\omega_0) = 1. \tag{3.18}$$

The upper bound is the same as in the previous case without damping ($\eta = 0$). It remains to compute the lower bound $a_j - b_j = 1 - b_j\big(1 + T_j(\omega_0)\big)$. By using

$$\lim_{j\to\infty} T_j\Big(1 + \frac{\eta}{m^2}\Big) = \cos(\sqrt{-2\eta}) = \cosh(\sqrt{2\eta})$$

(see [7, Chapter IV.2, exercise 8]) and differentiating this relation twice with respect to η, we obtain

$$\lim_{j\to\infty} a_j - b_j = \frac{\sinh(\sqrt{2\eta})\frac{1}{\sqrt{2\eta}} - 1}{2\sinh^2(\frac{\sqrt{2\eta}}{2})} \cong 0.32995. \tag{3.19}$$

Numerical computations for stage numbers ranging from 3 to 10000 show that $a_j - b_j$ is monotonically decreasing and confirm the lower bound (3.19).

Now for each corner of the square $[0.32995,\, 1]^2$, we calculate the values of $\tau\mu$ for which $\big|\widetilde{R}_{m+1}\big|^2 \le 1$.

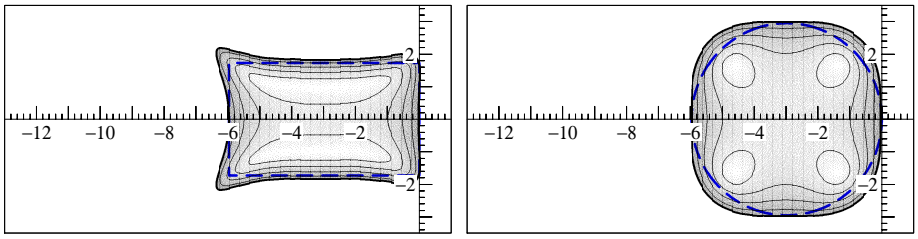

FIG. 3.5. *Stability regions in the $(\tau\lambda,\, \tau\mu)$-space of a 4-stage PRKC with $v = 1$ and $\alpha_3 = 0$ on the left and RKC on the right when $\eta \to \infty$. The dashed lines represent geometric shapes embedded inside these stability regions.*

Case $R_m = R_{m-1} = 1$: This case is identical to the undamped case (3.14).
Case $R_m = 0.32995,\ R_{m-1} = 1$: We have $\left|\widetilde{R}_{m+1}\right|^2 \le 1$ for $\tau\mu \in [-1.7273,\ 1.7273]$.
Case $R_m = R_{m-1} = 0.32995$: Since $c_{m-1} \ge 0.379$ (which is true for $m \ge 3$, see (2.5)), we get the restriction $\tau\mu \in \left[-1.7714,\ 1.7714\right]$.
Case $R_m = 1,\ R_{m-1} = 0.32995$: By using $c_{m-1} \ge 0.3$ for $m \ge 3$, we obtain $\tau\mu \in \left[-2.0333,\ 2.0333\right]$.

For $m = 2$, third and fourth cases follow from the remark at the end of the proof of the Theorem 3.2 and $c_{m-1} = 13/54$. ∎

For several values of m, we set $k_1 = \beta_{\mathbb{R}}(m)$ and computed k_2 such that the rectangle defined by $\{(x,\, y) \in \mathbb{R}^2;\ -k_1 \le x \le 0,\ |y| \le k_2\}$ is included in the stability region, see Table 3.1. Numerical results give values for k_2 slightly higher than those proved in Theorem 3.3.

TABLE 3.1
Comparison of constants k_1 and k_2 such that $\{(x,\, y) \in \mathbb{R}^2;\ -k_1 \le x \le 0,\ |y| \le k_2\}$ is included in the stability region of m-stage PRKC with damping parameters $\eta = 2/13$ and ∞.

$\backslash m$		4	6	10	20	200	1000
$\eta = 2/13$	k_1	9.75	22.75	64.35	259.35	25999.35	649999.35
	k_2	1.73	1.73	1.73	1.73	1.73	1.73
$\eta = \infty$	k_1	6	10	18	38	398	1998
	k_2	1.73	1.73	1.72	1.68	1.64	1.64

3.4. Infinite damping. For the infinitely damped case $\eta \to \infty$ (only for $m \ge 3$ because $R_m(z)$ does not depend on η for $m = 2$), by using

$$T_j(x) \sim 2^{j-1}x^j, \quad T_j'(x) \sim j2^{j-1}x^{j-1}, \quad T_j''(x) \sim j(j-1)2^{j-1}x^{j-2}$$

for $x \gg 1$, we substitute

$$R_{m-1}(\tau\lambda) \sim \frac{1}{m-1} + \frac{m-2}{m-1}\Big(1 + \frac{\tau\lambda}{m-1}\Big)^{m-1},$$
$$R_m(\tau\lambda) \sim \frac{1}{m} + \frac{m-1}{m}\Big(1 + \frac{\tau\lambda}{m-1}\Big)^{m}, \tag{3.20}$$

and $c_{m-1} = (m-2)/(m-1)$ into $\widetilde{R}_{m+1}(\tau\lambda,\, \tau\mu)$ in (3.16). As before we computed k_1 and k_2 for several values of m such that the rectangle defined by $\{(x,\, y) \in \mathbb{R}^2;\ -k_1 \le x \le 0,\ |y| \le k_2\}$ is included in the stability region, see Table 3.1. We observed

numerically that in the infinitely damped case k_2 decreases slowly and monotonically from 1.73 to 1.64 for m ranging from 3 to 1000. In contrast, the stability region of RKC (2.2) becomes circular (for $\eta \to \infty$) with center point $1 - m$ and radius $m - 1$ (illustrated in FIG 3.5). Furthermore, as the quadratic increase of the real stability boundary $\beta_{\mathbb{R}}(m)$ with the number of stages m for a small damping η turns linear for large η, *viz.* the real stability boundary in FIG 3.5 becomes $\beta_{\mathbb{R}}(m) \approx 2(m-1)$, there is no advantage to take a large damping for this configuration.

4. Advection-diffusion-reaction problems. A main question for stabilized methods applied to an advection-diffusion problem is how to choose the number of stages m and the step size τ. In contrast to classical RKC methods [23, 17], the choice of the damping parameter η is not very critical for our method (see the discussion of Sect. 3.4).

We consider the n-dimensional scalar model

$$u_t + a \cdot \nabla u = d\Delta u, \tag{4.1}$$

where $a = (a_1, \ldots, a_n)$ and d are taken constant, and the partial derivatives u_{x_k} and $u_{x_k x_k}$ are discretized on uniform grids with mesh width h_k with the second-order central scheme.

4.1. The pure advection or diffusion case. For the pure diffusion case ($G \equiv 0$), we operate in the same way as in [8, 22, 23], because the F-method is RKC. For any given step size τ, we choose the minimal stage number m satisfying the stability condition $\tau\sigma \le \beta_{\mathbb{R}}(m)$, where $\sigma \le 4d \sum_{k=1}^{n} {h_k}^{-2}$ denotes the spectral radius of $\frac{\partial F}{\partial Y}(t, Y)$. We obtain the step size restriction

$$\tau \le \frac{1}{2d \sum_{k=1}^{n} {h_k}^{-2}} \cdot \frac{\beta_{\mathbb{R}}(m)}{2} \tag{4.2}$$

For the pure advection case ($F \equiv 0$), the G-method shares its stability function with any one-step 3-stage explicit Runge-Kutta method of order 3, *i.e.*

$$R(z) = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!}.$$

For the central difference scheme the Jacobian matrix of the advection term G is diagonalisable with eigenvalues $i\mu$ bounded by $|\mu| \le \sum_{k=1}^{n} |a_k| \, {h_k}^{-1}$. Since the PRKC method has the imaginary stability boundary $\beta_{\mathbb{I}} = \sqrt{3}$ (see Sect. 3.3), we restrict the step size according to

$$\tau \le \frac{1.7}{\sum_{k=1}^{n} |a_k| \, {h_k}^{-1}}. \tag{4.3}$$

4.2. The advection-diffusion case. The embedded rectangles $\mathcal{R}(m)$ inside the stability region will facilitate the selections of the step size τ and the number of stages m. In fact for (4.1), Wesseling's geometric approach [24, Chapter 5] gives two step size restrictions which guarantee that the eigenvalues emerging from von Neumann stability analysis lie inside a rectangle [24, Theorem 5.7.1]. These τ-restrictions are nothing else than (4.2) and (4.3). Then for advection-diffusion problems, we just combine the restrictions as follows:

For a given trial step size τ (obtained from a local error estimation procedure),

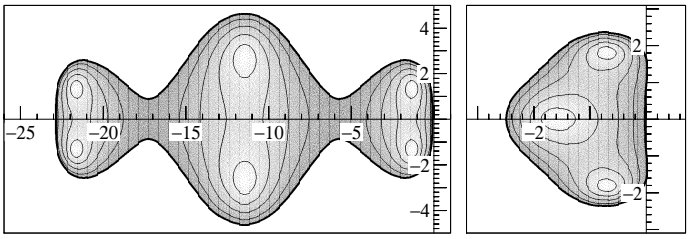

FIG. 4.1. *Stability regions for the* 6*-stage damped RKC method on the left and the 3-stage explicit RK method of order* 3 *on the right.*

1. If the condition (4.3) is not satisfied, reduce τ such that $\tau \sum_{k=1}^{n} |a_k|\, {h_k}^{-1} = 1.7$.
2. Compute the number of stages m to satisfy (4.2).

This algorithm, which is justified for problems of the form (4.1), will also be used in more general situations. For example, for nonlinear advection-diffusion systems such as $u_t + \nabla \cdot (\underline{a} u) = \nabla \cdot (D \nabla u)$ with $\underline{a} = \underline{a}(u)$ and $D = D(u)$ (positive diagonal matrix), we insert the maximum values of the velocities a and the diffusion coefficient d as in [23].

4.3. Additional reaction terms. Unlike previous cases there are two ways to treat mildly stiff or non-stiff reaction terms of an advection-diffusion-reaction problem, either with the F-method or the G-method. Recall that the eigenvalues of the diffusion and the non-stiff reaction terms are proportional to $\mathcal{O}(h^{-2})$ and $\mathcal{O}(1)$, respectively (here h denotes the mesh width of the spatial discretization). At each step those eigenvalues are multiplied by the step size τ which is determined to keep the eigenvalues of the diffusion term inside the stability region.

As the F-method is RKC, we chose the damping $\eta = 2/13$ to obtain a more robust method (see the discussion at the end of Sect. 2). If the reaction term is treated by the F-method, it can be regarded as a small perturbation of the diffusion and should not cause instability. The advantage of this approach is that mildly stiff reaction can be treated, its disadvantage is that the number of function evaluations of the reaction term is equal to that of the diffusion.

The second alternative with the G-method is more suitable when the reaction is non-stiff, because each step only asks for 4 function evaluations of the reaction term. FIG 4.1 illustrates separately the stability regions for the damped RKC (F-method) and the 3-stage explicit RK of order 3 (G-method). Note that the stability region on the right of FIG 4.1 has its imaginary stability boundary equal to $\sqrt{3}$, while for any two-stage explicit RK of order 2, $\beta_{\mathbb{I}}$ is equal to 0.

5. Software issues. The FORTRAN95 program `PRKC` is a variable step size, variable formula code that uses explicit Runge-Kutta formulas to solve efficiently a class of large systems of mildly stiff ODEs. It is based on the codes [1, 15, 14]. We will only describe the new aspects of the code. For further information about codes based on an explicit Runge-Kutta formula, we refer the reader to [6, 7].

5.1. Estimate of the local error. The main innovation is the local error estimate used in selecting the step size. The program uses two different estimates and it takes the maximum value to select the new step size following standard strategies, see *e.g.* [6, p. 167]. Then it applies the algorithm described in the next Subsection 5.2 to select the number of stages m.

F-error estimate. Observe that if $\alpha_1 = \alpha_2 = \alpha_3 = 0$ in (3.1), then the internal stages K_0 to K_m are identical to the internal stages of the RKC method with initial approximation $Y_0 + \alpha_0 \tau G_1$. Thus, before adding the α-terms in K_m, we calculate a local error estimate as proposed in [15]. Note that this estimate requires one additional evaluation of $F(t, Y)$ (compared to the RKC method).

G-error estimate. The second error estimate is given by the difference to the numerical approximation of an embedded RK formula (3.1) with coefficients $\widehat{\alpha}_i$ in place of α_i. We always assume $\widehat{\alpha}_0 = \alpha_0 = \frac{1}{2}$ so that the first m stages are identical for both methods. If the coefficients satisfy

$$\widehat{\alpha}_0 + \widehat{\alpha}_1 + \widehat{\alpha}_2 + \widehat{\alpha}_3 = 1, \quad \widehat{\alpha}_0(\widehat{\alpha}_2 + \widehat{\alpha}_3) = 1/2, \quad \widehat{\alpha}_0 = 1/2, \quad c_{m-1}\widehat{\alpha}_3 = 1/2, \quad (5.1)$$

then the approximation $\widehat{Y}_{n+1} = \widehat{K}_m$ has order 2. By setting

$$\widehat{\alpha}_1 = -1/2, \quad \widehat{\alpha}_2 = 1 - \frac{1}{2c_{m-1}}, \, \widehat{\alpha}_3 = \frac{1}{2c_{m-1}}, \quad (5.2)$$

we thus obtain an embedded RK method that has one internal stage less than the PRKC method. For $F \equiv 0$, the embedded method becomes the explicit midpoint rule

$$\begin{array}{c|cc} 0 & & \\ 1/2 & 1/2 & \\ \hline & 0 & 1 \end{array}$$

Note that other embedded methods are possible. For instance, for $v = 1$ and putting $\widehat{\alpha}_i = \alpha_i$ for $i = 1, 2, 3$, $\widehat{\alpha}_4 = \widehat{\alpha}_5 = \widehat{\alpha}_6 = 0$ and $\widehat{\alpha}_7 = 1/2$ the approximation $\widehat{Y}_{n+1} = \widehat{K}_{m+1}$ can be considered as the trapezoidal rule.

5.2. Control of step-size and number of stages. `PRKC` requires several initialization parameters, for instance the initial step size guess τ_0, the absolute and relative error tolerances *Atol* and *Rtol*, the functions F and G and their spectral radius estimates σ_F and σ_G, respectively.

An ideal problem for `PRKC` is a problem giving rise to a system $\dot{y} = Ay + By$ where A and B can be simultaneously diagonalized and the eigenvalues of A are negative real, and those of B on the imaginary axis. For example, the problem $u_t + au_x = du_{xx}$ with spatial periodicity and space discretization by second-order central differences. For this kind of problems the modified test equation (3.9) is particularly well suited, and then σ_F and σ_G can be identified with λ and μ, respectively.

Unfortunately most systems $\dot{y} = Ay + By$, resulting from a space discretization with second-order schemes, do not have the properties of the ideal problem. The eigenvalues of A are close to the negative real axis, and those of B may be close to the imaginary axis or not. If the eigenvalues of the Jacobian of G are known to lie close to the imaginary axis, we recommend to set σ_G as being their maximum modulus, otherwise we set $\sigma_G := 0$ to disable the advection restriction (4.3) (on the fly mode).

Now we suppose that a trial step size τ, obtained from the initial step size guess or from the local error estimates described in Subsection 5.1, is given. Based on the equations (4.2) and (4.3), `PRKC` performs the following algorithm :

1. If $\sigma_G \tau > 1.7$, then it reduces τ such that $\sigma_G \tau = 1.7$.
2. It computes the number of stages m by

$$\sigma_F \tau = \beta_{\mathbb{R}}(m) \approx 0.65(m^2 - 1)$$

The first and second steps can be interpreted as an advection and a diffusion restriction, respectively. On the fly mode ($\sigma_G = 0$) only considers diffusion restriction, while the normal mode ($\sigma_G > 0$) considers both.

6. Numerical examples. We will present numerical results obtained for three test problems. In Section 6.1, a constant coefficient 1D type advection-diffusion problem is solved and in Section 6.2, we deal with a nonlinear 1D type diffusion-reaction problem. The third problem is a parabolic initial-boundary value problem which involves nonlocal integral terms over the spatial domain. The code and the drivers are publicly available on the page http://www.unige.ch/∼zbindech/.

6.1. Advection-diffusion 1D. We consider the periodic advection-diffusion problem in 1-dimension

$$\left\{ \begin{array}{rcl} u_t + a u_x & = & d u_{xx} \\ u(x,0) & = & u_0(x) \\ u(1,t) & = & u(0,t) \end{array} \right. \tag{6.1}$$

with constant coefficients $a \in \mathbb{R}$, $d > 0$, initial condition $u_0 \in \mathrm{L}_2[0,1]$ over the space $x \in [0,1]$ and time $t \geq 0$. For the space discretization on a uniform grid $\{x_1, x_2, \ldots, x_N\}$ with grid points $x_j = jh$ and mesh width $h = 1/N$, we use second-order central differences for the advection and diffusion terms. We obtain the semi-discrete system

$$w_j'(t) = \left(\frac{d}{h^2} + \frac{a}{2h} \right) w_{j-1}(t) - \frac{2d}{h^2} w_j(t) + \left(\frac{d}{h^2} - \frac{a}{2h} \right) w_{j+1}(t), \tag{6.2}$$

where $j = 1, \ldots, N$ and $w_0(t) = w_N(t)$, $w_{N+1}(t) = w_1(t)$. Fourier analysis gives the eigenvalues

$$\lambda_k = \frac{2d}{h^2} \left(\cos(2\pi k h) - 1 \right) - \frac{\mathrm{i}a}{h} \sin(2\pi k h), \quad k = 1, \ldots, N. \tag{6.3}$$

These eigenvalues are located on an ellipse in the left half-plane $\mathbb{C}^-$. The discrete Fourier transform of the vector $w(0) \in \mathbb{R}^N$ (initial condition of the semi-discrete system) is

$$w(0) = \sum_{k=1}^{N} z_k \phi_k \qquad \text{with} \qquad z_k = \frac{1}{N} \sum_{j=1}^{N} u_0(x_j) (\overline{\phi_k})_j$$

where $\phi_k = \left(e^{2\pi \mathrm{i} k x_1}, e^{2\pi \mathrm{i} k x_2}, \ldots, e^{2\pi \mathrm{i} k x_N} \right)^T \in \mathbb{C}^N$ for $k = 1, \ldots, N$ denote the discrete Fourier modes. The solution $w(t)$ of the semi-discrete system (6.2) arising from the problem (6.1) is given by

$$w(t) = \sum_{k=1}^{N} z_k e^{\lambda_k t} \phi_k. \tag{6.4}$$

The problem (6.1) gives us an excellent benchmark to compare the performance of our proposed program `PRKC` with the program `RKC` [15]. We take $a = 0.1$, $d = 1$, $u_0(x) = \sin(2\pi x)$ and $N = 64,\ 128$. For numerical integration with `PRKC` we define the functions

$$F(u) = d u_{xx}, \qquad G(u) = -a u_x$$

TABLE 6.1
Comparaison of PRKC and RKC on the advection-diffusion problem (6.1). *The numbers to the left of a slash correspond to PRKC, those to the right correspond to RKC.*

Tol	Steps	F-evals	G-evals	$m_{average}$	m_{max}	L_∞-error
$h = 1/64$						
10^{-1}	7/5	128/109	28/109	17/21	31/32	$1.7 \times 10^{-2}/1.7 \times 10^{-2}$
10^{-2}	9/8	154/139	36/139	16/17	24/26	$3.5 \times 10^{-2}/4.3 \times 10^{-3}$
10^{-3}	13/14	194/189	52/189	14/13	19/19	$9.9 \times 10^{-4}/9.1 \times 10^{-4}$
10^{-4}	24/27	280/268	96/268	11/10	14/14	$2.1 \times 10^{-4}/2.0 \times 10^{-4}$
10^{-5}	48/55	421/397	192/397	8/7	11/10	$5.3 \times 10^{-5}/4.2 \times 10^{-5}$
$h = 1/128$						
10^{-1}	7/5	243/213	28/213	34/42	62/63	$1.7 \times 10^{-2}/1.7 \times 10^{-2}$
10^{-2}	9/8	293/269	36/269	32/33	48/50	$3.5 \times 10^{-3}/4.2 \times 10^{-3}$
10^{-3}	13/14	369/366	52/366	27/26	37/38	$9.7 \times 10^{-4}/9.0 \times 10^{-4}$
10^{-4}	24/27	523/519	96/519	21/19	24/27	$2.1 \times 10^{-4}/2.0 \times 10^{-4}$
10^{-5}	48/54	762/750	192/750	15/14	20/19	$4.8 \times 10^{-5}/4.2 \times 10^{-5}$

(in fact $F(u)$ and $G(u)$ describe the semi-discrete system (6.2)). We set the initial guess for the step size to $\tau_0 = 10^{-3}$ for PRKC to avoid rejection of the first step. For the approximation of the spectral radii $\frac{\partial F}{\partial Y}(t, Y)$ and $\frac{\partial G}{\partial Y}(t, Y)$, we respectively take the maximum modulus of eigenvalues emerging from von Neumann stability analysis of the system $u_t = du_{xx}$ and $u_t = -au_x$, *i.e.* $\sigma_F = 4dN^2$ and $\sigma_G = |a|\, N$. We integrate over $0 < t \leq 1/10$ with values of $Tol = Atol = Rtol$ ranging from 10^{-1} to 10^{-5}. Table 6.1 presents the numerical results obtained with variable step sizes, L_∞-error denotes the error at $t = 1/10$ with respect to the solution (6.4).

As the space discretization is of second order $w_j(t) = u(x_j, t) + \mathcal{O}(h^2)$ with mesh width $h = 1/64,\ 1/128$, it is not necessary to take a very fine tolerance. Note that if h is divided by 2 then the spectral radius estimate σ_F is multiplied by 2^2, but the average number of stages is only multiplied by 2 because of the quadratic growth of the real stability interval with respect to the number of stages that keeps the same number of steps. This implies that the number of evaluations of F is multiplied by 2 while the number of evaluations of G remains the same for PRKC.

6.2. Brusselator 1D. We consider the Brusselator problem [6, 7] in one spatial variable x

$$\begin{aligned} u_t &= A + u^2 v - (B+1)u + \alpha u_{xx} \\ v_t &= Bu - u^2 v + \alpha v_{xx} \end{aligned} \tag{6.5}$$

with $x \in [0,\ 1]$, $A = 1$, $B = 3$, $\alpha = 1/50$ and boundary conditions

$$\begin{aligned} &u(0,t) = u(1,t) = 1, \qquad v(0,t) = v(1,t) = 3, \\ &u(x,0) = 1 + \sin(2\pi x), \quad v(x,0) = 3. \end{aligned}$$

We replace the second spatial derivatives by finite differences on a grid of N points $x_i = i/(N+1)$ for $i = 1, \ldots, N$. We define the functions as follows

$$F(u,v) = \begin{pmatrix} \alpha u_{xx} \\ \alpha v_{xx} \end{pmatrix}, \qquad G(u,v) = \begin{pmatrix} A + u^2 v - (B+1)u \\ Bu - u^2 v \end{pmatrix}.$$

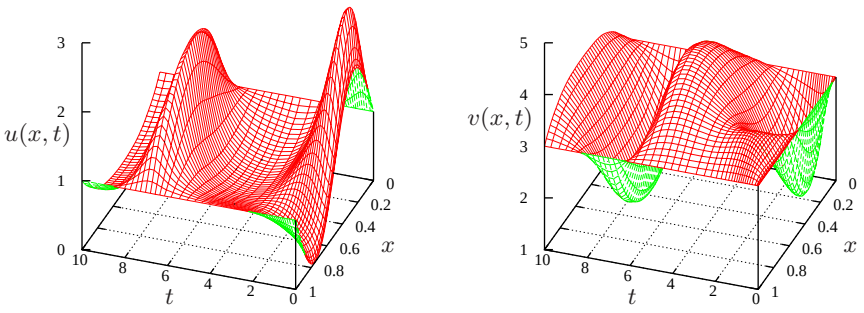

FIG. 6.1. *Solutions $u(x,t)$ and $v(x,t)$ of Brusselator 1D with* `PRKC`.

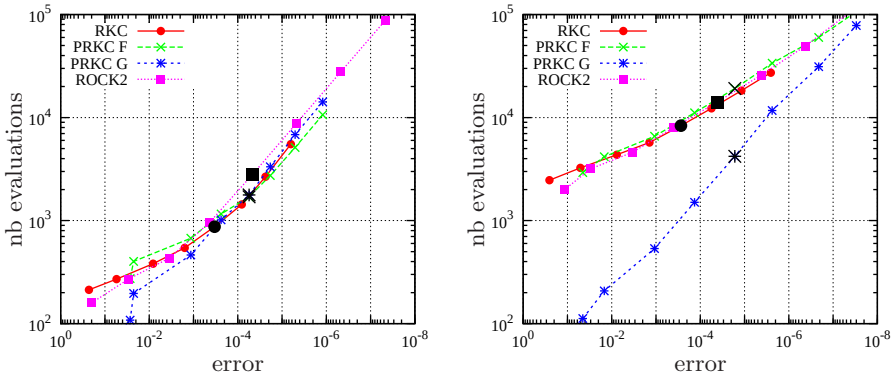

FIG. 6.2. *Comparison of Chebyshev codes on Brusselator 1D: left $N = 40$, right $N = 500$.*

We integrate over $0 \le t \le 10$ with $Atol = Rtol = Tol = 10^{-n}$ for $n = 1, 2, \ldots, 8$ with three codes; `RKC` [15], `PRKC` (on the fly mode, *i.e.* without the advection restriction (4.3)) and `ROCK2` [4]. FIG 6.2 shows the number of function evaluations *versus* the L_∞-error at $t = 10$. The tolerance 10^{-5} is distinguished by an enlarged symbol. We observe that for a grid of 40 points, our method is not more attractive than `RKC` or `ROCK2`. The reason is that it is not necessary to use a large number of stages to solve the problem. However, when the problem is very stiff ($N = 500$), `PRKC` reduces significantly the number of function calls of the non-stiff terms.

6.3. Parabolic integro-differential equation 1D. Following [20], we consider the simplified model of the temperature profile of air near the ground

$$\left\{ \begin{array}{rcl} u_t & = & u_{xx} + \sigma \displaystyle\int_0^1 f\left(s,t,u(s,t)\right) k\left(|x-s|\right) ds \\ u(x,0) & = & \xi\left(\cos(\pi x)+1\right) \\ u(0,t) & = & \xi\left(2-\sqrt{t}\right) \\ u_x(1,t) & = & 0 \end{array} \right. \tag{6.6}$$

where $f(x,t,u) = -u^4$ and $k(y) = 1/(1+y)^2$ with $x \in [0,1]$, $0 \le t \le 1$ and $\sigma \ge 0$, $\xi \ne 0$ are parameters. For the space discretization we use a uniform mesh with

$N = 100$ mesh intervals. We approximate the Laplacian term by the second-order central differences and the integral term by the second-order composite trapezoidal rule to obtain a semi-discrete system in $w(t)$ where $w_j(t) \approx u(x_j, t)$. For the numerical integration with PRKC, the F-function corresponds to the discretization of u_{xx} and the G-function to the discretization of the integral. In particular, the i-th component of the G-function at time t is given by

$$G_i(t,\, w(t)) = \frac{\sigma}{N} \sum_{j=0}^{N}{}' f(x_j, t, w_j(t)) k(|x_i - x_j|),$$

where the prime indicates that the first and last terms of the sum are divided by 2. Therefore an evaluation of the G-function is of order N^2 operations, and much more expensive than that of the F-function.

TABLE 6.2

Comparaison of PRKC and RKC on the parabolic integro-differential equation (6.6). *The numbers to the left of a slash correspond to PRKC, those to the right correspond to RKC.*

σ	Steps	F-evals	G-evals	$m_{average}$	m_{max}	L_∞-error
$N = 100$, $Tol = 10^{-3}$						
1	571/25	6918/1026	2284/1026	11/44	16/96	$3.4 \times 10^{-4}/9.1 \times 10^{-4}$
10^{-1}	162/25	3384/1025	648/1025	20/44	25/96	$3.3 \times 10^{-3}/4.3 \times 10^{-3}$
10^{-2}	51/25	1760/1026	204/1026	34/44	44/96	$7.5 \times 10^{-5}/1.0 \times 10^{-3}$
10^{-3}	26/25	1101/1026	104/1026	41/44	78/96	$5.8 \times 10^{-4}/1.1 \times 10^{-3}$

We integrate over $0 \le t \le 1$ with $Atol = Rtol = Tol = 10^{-3}$, $\xi = 1/2$ and $\sigma = 10^{-n}$ for $n = 0, 1, 2, 3$ with RKC [15] ($\sigma_F = 4N^2$) and PRKC ($\sigma_F = 4N^2$, $\sigma_G = 0$ and initial guess $\tau_0 = 10^{-3}$). Table 6.2 presents the numerical results obtained with variable step sizes, L_∞-error denotes the error at $t = 1$ with respect to reference solutions computed with Radau5 [7]. Our method gives excellent results for small values of ξ and σ in which case the Lipschitz constant of the G-function is not large. In other cases, one can use PRKC as RKC by defining the whole problem in the F-function and putting $G \equiv 0$.

Acknowledgments. The author is grateful to Ernst Hairer for the useful discussions and comments, and acknowledges the support from the Swiss National Science Foundation, project No. 121561/1. The author would also like to thank the anonymous referees for their constructive comments.

REFERENCES

[1] A. Abdulle. Fourth order Chebyshev methods with recurrence relation. *SIAM J. Sci. Comput.*, 23(6):2041–2054 (electronic), 2002.

[2] A. Abdulle and S. Cirilli. S-ROCK: Chebyshev methods for stiff stochastic differential equations. *SIAM J. Sci. Comput.*, 30(2):997–1014, 2008.

[3] A. Abdulle, Y. Hu, and T. Li. Chebyshev methods with discrete noise: the τ-ROCK methods. *J. Comput. Math.*, 28(2):195–217, 2010.

[4] A. Abdulle and A. A. Medovikov. Second order Chebyshev methods based on orthogonal polynomials. *Numer. Math.*, 90(1):1–18, 2001.

[5] M. P. Calvo, J. de Frutos, and J. Novo. Linearly implicit Runge-Kutta methods for advection-reaction-diffusion equations. *Appl. Numer. Math.*, 37(4):535–549, 2001.

[6] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving ordinary differential equations. I*, volume 8 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 1993. Nonstiff problems.

[7] E. Hairer and G. Wanner. *Solving ordinary differential equations. II*, volume 14 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 1996. Stiff and differential-algebraic problems.

[8] W. Hundsdorfer and J. G. Verwer. *Numerical solution of time-dependent advection-diffusion-reaction equations*, volume 33 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2003.

[9] C. A. Kennedy and M. H. Carpenter. Additive Runge-Kutta schemes for convection-diffusion-reaction equations. *Appl. Numer. Math.*, 44(1-2):139–181, 2003.

[10] V. I. Lebedev. How to solve stiff systems of differential equations by explicit methods. In *Numerical methods and applications*, pages 45–80. CRC, Boca Raton, FL, 1994.

[11] V. I. Lebedev. Explicit difference schemes for solving stiff problems with a complex or separable spectrum. *Zh. Vychisl. Mat. Mat. Fiz.*, 40(12):1801–1812, 2000.

[12] J. Martín-Vaquero and B. Janssen. Second-order stabilized explicit Runge-Kutta methods for stiff problems. *Computer Physics Communications*, 180(10):1802–1810, 2009.

[13] A. A. Medovikov. High order explicit methods for parabolic equations. *BIT*, 38(2):372–390, 1998.

[14] L. F. Shampine, B. P. Sommeijer, and J. G. Verwer. IRKC: an IMEX solver for stiff diffusion-reaction PDEs. *J. Comput. Appl. Math.*, 196(2):485–497, 2006.

[15] B. P. Sommeijer, L. F. Shampine, and J. G. Verwer. RKC: an explicit solver for parabolic PDEs. *J. Comput. Appl. Math.*, 88(2):315–326, 1998.

[16] B. P. Sommeijer and J. G. Verwer. *A performance evaluation of a class of Runge-Kutta-Chebyshev methods for solving semi-discrete parabolic differential equations*. Afdeling Numerieke Wiskunde [Department of Numerical Mathematics], 91. Mathematisch Centrum, Amsterdam, 1980.

[17] B. P. Sommeijer and J. G. Verwer. On stabilized integration for time-dependent PDEs. *J. Comput. Phys.*, 224(1):3–16, 2007.

[18] M. Torrilhon and R. Jeltsch. Essentially optimal explicit Runge-Kutta methods with application to hyperbolic-parabolic equations. *Numer. Math.*, 106(2):303–334, 2007.

[19] P. J. van der Houwen and B. P. Sommeijer. On the internal stability of explicit, m-stage Runge-Kutta methods for large m-values. *Z. Angew. Math. Mech.*, 60(10):479–485, 1980.

[20] A. S. Vasudeva Murthy and J. G. Verwer. Solving parabolic integro-differential equations by an explicit integration method. *J. Comput. Appl. Math.*, 39(1):121–132, 1992.

[21] J. G. Verwer, W. H. Hundsdorfer, and B. P. Sommeijer. Convergence properties of the Runge-Kutta-Chebyshev method. *Numer. Math.*, 57(2):157–178, 1990.

[22] J. G. Verwer and B. P. Sommeijer. An implicit-explicit Runge-Kutta-Chebyshev scheme for diffusion-reaction equations. *SIAM J. Sci. Comput.*, 25(5):1824–1835 (electronic), 2004.

[23] J. G. Verwer, B. P. Sommeijer, and W. Hundsdorfer. RKC time-stepping for advection-diffusion-reaction problems. *J. Comput. Phys.*, 201(1):61–79, 2004.

[24] P. Wesseling. *Principles of computational fluid dynamics*, volume 29 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2001.